

X. L'analisi della dinamica del network con SIENA[♦]

di Laura Savoia^{*}

SIENA, acronimo di *Simulation Investigation for Empirical Network Analysis*, è un programma per la stima di modelli sull'evoluzione nel tempo di un network ed è inserito nel sistema StOCNET (Boer, Huisman, Snijders, Zeggelink 2005) per l'analisi statistica avanzata delle reti sociali, sviluppato da un gruppo di ricercatori olandesi coordinati dal professor Tom Snijders della Rijksuniversiteit Groningen (NL). Il programma si può scaricare gratuitamente all'indirizzo web <http://stat.gamma.rug.nl/stocnet/>. In internet si possono trovare anche un gruppo di discussione per StOCNET all'indirizzo <http://groups.yahoo.com/group/stocnet/> e il sito di Snijders, all'indirizzo <http://stat.gamma.rug.nl/snijders/>, con una sezione specifica dedicata a SIENA. Nel sito sono disponibili anche alcuni *papers* che approfondiscono e spiegano dettagliatamente la logica e la matematica sviluppata in StOCNET.

La presente elaborazione intende fornire un'introduzione alla logica del modello e all'utilizzo di SIENA. Il primo paragrafo illustra il funzionamento del modello e si sofferma sugli aspetti tecnici connessi alle due funzioni che lo definiscono (*objective function* e *rate function*). Il secondo paragrafo ripercorre passo per passo le procedure da seguire nel programma per l'elaborazione del modello, illustrando i passaggi operativi su un esempio concreto relativo allo scambio di giocatori tra le squadre della massima serie di pallamano in Italia, dall'inserimento dei dati all'output prodotto.

[♦] Questo articolo sarà pubblicato all'interno del libro *Analisi delle reti sociali. Teorie, metodi, applicazioni* a cura di Andrea Salvini, edito da FrancoAngeli (2007).

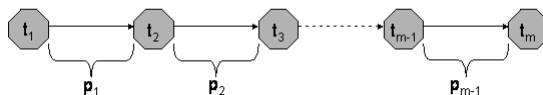
^{*} Università degli Studi di Trento, laura.savoia@soc.unitn.it

1. Il modello dinamico *actor-oriented* per l'evoluzione del network

Le relazioni tra gli attori sociali non sono immutabili nel tempo; anche i legami più stabili hanno un inizio e una fine. Osservando la rete o configurazione di una determinata relazione all'interno di uno stesso gruppo in almeno due momenti temporali successivi si esamina l'evoluzione o la dinamica di tale network. Questa evoluzione si può modellare e spiegare come funzione di effetti strutturali e/o delle caratteristiche individuali degli attori (Snijders, Steglich, Schweinberger 2005).

Gli effetti strutturali sono meccanismi endogeni del network che si collegano alla probabilità dei legami tra gli attori; ad esempio, in una *community* su internet è ragionevole ipotizzare una tendenza alla reciprocazione dei legami, per cui è più probabile che un soggetto A scriva una mail ad un soggetto B in un momento t_m nel caso in cui in un momento t_{m-1} B abbia scritto ad A che nel caso in cui queste due persone non si siano mai scritte prima. Anche gli attributi specifici dei soggetti possono favorire o inibire un legame tra due attori, sulla base dei meccanismi di omofilia o eterofilia¹ che governano i diversi tipi di relazione; ad esempio è più probabile che un tifoso del Milan vada a vedere una partita di calcio accompagnato da un altro milanista che da uno juventino o un interista (omofilia), mentre nella maggior parte dei casi ballano insieme un uomo e una donna e non due persone dello stesso sesso (eterofilia).

Fig. 1 - Catena markoviana con parametro. a tempo continuo osservato in momenti discreti



In pratica, si costruisce uno schema evolutivo che aiuti a comprendere l'effettiva dinamica del network. Dietro alla logica di SIENA c'è un modello basato su catene markoviane con un parametro a tempo continuo osservato in momenti temporali discreti². Nello studio delle reti di relazioni tale schema matematico si applica "fotografando" il network in momenti successivi e partendo dal presupposto che negli intervalli tra i momenti di osservazione abbia luogo una continua evoluzione non osservabile. La figura 1 offre una

¹ Per una rassegna dettagliata della letteratura scientifica sull'omofilia nelle reti di relazioni sociali cfr. McPherson, Smith-Lovin, Cook (2001).

² Per una trattazione dettagliata di questi processi stocastici cfr., ad esempio, Norris (1997) oppure Taylor, Karlin (1998).

rappresentazione grafica di questa situazione: si hanno m momenti di osservazione ($t_1, t_2, t_3, \dots, t_{m-1}, t_m$) del network e tra questi momenti si situano $m-1$ periodi o intervalli (p_1, p_2, p_{m-1}) durante i quali il network si evolve continuamente ma non è osservato.

Da quando Holland e Leinhardt (1977a, 1977b) e Wasserman (1977) hanno proposto l'applicazione di schemi interpretativi di questo tipo ai fenomeni sociali per lo studio di dati relazionali longitudinali, gli studiosi hanno elaborato diversi modelli, da quello della reciprocità proposto da Wasserman (1977, 1979, 1980) e integrato da Leenders (1995, 1996) con la similarità a quello della popolarità, sempre di Wasserman (1977, 1980). Snijders (1996, 2001, 2005a) ha sviluppato un modello che prende in considerazione contemporaneamente più effetti strutturali del network e interpreta l'evoluzione della rete di relazioni tra gli attori di un network come conseguenza delle scelte degli attori stessi di creare nuovi legami oppure interrompere quelli esistenti all'interno del network. Si tratta di relazioni dirette, pertanto ogni attore può controllare solo i legami in uscita (rappresentati nel relativo vettore riga della matrice di adiacenza), ad esempio in una rete di posta elettronica si può decidere se scrivere a qualcuno ma non se ricevere posta.

Essendo costruito su processi markoviani, il modello assume che gli attori operino le proprie scelte a partire dalle informazioni che hanno sulla struttura relazionale interna al network, mentre le strutture che la rete ha assunto in passato non sono rilevanti. In pratica per conoscere la configurazione della rete al momento t_m ci si basa esclusivamente sulle informazioni relative al momento t_{m-1} , che influenza le scelte degli attori e con esse la configurazione futura. Il momento iniziale di osservazione è il punto di partenza, su cui si condiziona l'evoluzione, pertanto non si modella. I cambiamenti nel network si verificano *step by step*, cioè un solo attore alla volta effettua un solo cambiamento nella propria struttura relazionale. Questi *microsteps* si accumulano e danno luogo ad ampi cambiamenti.

Nel modello gli attori agiscono nel tentativo di ottimizzare la configurazione del network in relazione all'utilità e ai propri obiettivi (Huisman, Snijders 2003). Il modello assume che tutti gli attori abbiano a disposizione un quadro completo della struttura relazionale interna al network; sulla base di tale informazione essi valutano la propria situazione all'interno della rete di relazioni e operano cambiamenti allo scopo di migliorare questa posizione. Gli attori sono indipendenti tra loro e agiscono senza coordinarsi, si preoccupano solo del breve periodo e non fanno supposizioni sulle eventuali contromosse di altri attori; in breve applicano una strategia miope. Le valutazioni in merito alla posizione vantaggiosa

all'interno del network sono funzione del punto di vista dell'attore (individualismo metodologico); in questo senso Snijders parla di modello orientato dall'attore, cioè *actor-oriented*.

Gli assunti del modello rappresentano certamente limitazioni rispetto alle situazioni reali e vanno presi con la dovuta cautela, per l'interpretazione delle diverse reti di relazioni sociali. Il modello non sostiene, ad esempio, che ci sia effettivamente una tendenza intenzionale e cosciente all'ottimizzazione da parte degli attori, però si può applicare il punto di vista dell'attore alle reti modellandole come se questa tendenza ci fosse. Un modello è sempre una semplificazione della realtà, un'approssimazione; è impossibile trovare una coincidenza perfetta tra schema interpretativo e realtà, anche perché non si possono prendere in considerazione tutte le forze in atto e non è possibile modellare l'effetto del caso.

Il modello *actor-oriented* si esprime attraverso due funzioni statistiche³. La prima, detta funzione di distribuzione nel tempo (*rate function*), è legata al momento in cui un attore cambia una sua relazione all'interno del network, specifica la frequenza con cui i singoli attori possono "fare una mossa" e può essere uguale per tutti oppure dipendere dagli attributi degli attori e/o dalla loro posizione all'interno del network. Il contenuto specifico della mossa che un attore effettua è determinato stocasticamente dalla seconda funzione del modello, detta funzione obiettivo (*objective function*). Tale funzione si suddivide in tre parti: una funzione di valutazione (*evaluation function*) che modella la soddisfazione degli attori per le diverse configurazioni possibili del network, una funzione di dotazione (*endowment function*⁴) che si collega alla gratificazione derivante dalle diverse mosse che hanno portato a tali configurazioni e una componente casuale riferita agli effetti residui.

Recentemente Snijders e i suoi collaboratori hanno sviluppato una versione più elaborata e complessa del modello *actor-oriented*, in grado di prendere in considerazione l'evoluzione congiunta della struttura del network e del comportamento degli attori che ne fanno parte (cfr. Steglich, Snijders, Pearson 2004; Snijders, Steglich, Schweinberger 2005). Spesso si può osservare che gli attori legati tra loro presentano attributi simili, ad esempio tra gli adolescenti è frequente che i membri di un gruppo di amici si vestano in modo

³ Nei testi più recenti sul modello *actor-oriented* risulta leggermente modificata la terminologia relativa a tali funzioni rispetto ai testi precedenti. Su indicazione personale di Snijders, qui è adottata questa nuova notazione.

⁴ In economia l'*endowment effect* (Thaler 1980) indica l'avversione alla perdita, determinata dal fatto che il valore che le persone attribuiscono ad un bene aumenta quando esse ne entrano in possesso; in termini micro-economici si tratta dello squilibrio tra prezzo di acquisto e di vendita.

simile, frequentino determinati locali, ascoltino la stessa musica, ecc. Per spiegare tali fenomeni di autocorrelazione bisogna tenere conto sia dei meccanismi di influenza o contagio (la relazione tra due soggetti influisce sulle loro scelte comportamentali: ascolto questa musica perché la ascoltano i miei amici) sia dei processi di selezione (un attributo individuale determina una scelta relazionale: ho fatto amicizia con Tizio perché frequentiamo lo stesso locale).

In questo nuovo modello gli attori, sempre nell'intento di massimizzare la propria posizione all'interno del network, possono modificare non solo le proprie relazioni ma anche il proprio comportamento: bisogna tenere in conto sia le decisioni relazionali sia quelle comportamentali e prevedere due funzioni obiettivo e due funzioni di attività, come si evince dalla tabella 1.

Tab. 1 - Modello per l'evoluzione congiunta del network e del comportamento

	<i>Contenuto del cambiamento</i>	<i>Frequenza del cambiamento</i>
<i>Decisioni relazionali</i>	Funzione obiettivo per le relazioni	Funz. di distribuz. nel tempo per le relazioni
<i>Decisioni comportamentali</i>	Funzione obiettivo per i comportamenti	Funz. di distribuz. nel tempo per i comportamenti

Fonte: adattamento da Steglich, Snijders, Pearson (2004)

In questo modo lo schema si complica da un punto di vista tecnico e operativo, tuttavia la logica del modello non cambia. La presente elaborazione si concentra proprio sull'aspetto logico e ricostruisce il modello *actor-oriented* per l'evoluzione del network. Per non appesantire la lettura, inoltre, si è scelto di non entrare nei dettagli matematici e statistici, per i quali si rimanda a Snijders (1996, 2001, 2005a, 2005b), Steglich, Snijders, Pearson (2004) e Snijders, Steglich, Schweinberger (2005).

1.1. Funzione obiettivo: funzione di valutazione (evaluation function)

La funzione di valutazione è il primo elemento della funzione obiettivo ed indica la soddisfazione che l'attore i prova per lo stato concreto x assunto dal network X . Si denota con l'espressione

$$f_i(\beta, x), \quad (1)$$

dove β è il vettore dei parametri che determinano la funzione e rappresenta l'intensità degli effetti strutturali e/o individuali presi in considerazione dal modello per determinare gli obiettivi e le preferenze dell'attore. In appendice 2 è riportata la tavola dei principali effetti per modellare l'evoluzione del

network, completa di formule statistiche e rappresentazioni grafiche esplicative.

Le principali forze strutturali che possono incidere sulle preferenze degli attori sono le seguenti:

- a. densità: propensione a creare legami arbitrari con un qualsiasi altro membro del network, questo effetto si misura attraverso l'*outdegree* dell'attore e costituisce una sorta di intercetta o effetto di controllo che va sempre inserito nel modello;
- b. reciprocità: inclinazione alla reciprocazione dei legami, cioè a creare un legame con un attore che è legato a noi;
- c. transitività: scelta di creare un legame con qualcuno, se questi ha legami con altri attori cui si è già legati, in parole povere si tratta della disposizione a fare amicizia con gli amici degli amici, disposizione che diventa più intensa all'aumentare del numero delle amicizie comuni;
- d. equilibrio: tendenza a creare legami con attori strutturalmente simili, cioè attori con cui si condividono altri legami, questo effetto si basa sul concetto di equivalenza strutturale;
- e. legami indiretti (o attori a distanza due): disposizione ad avere uno o più intermediari nel rapporto con gli altri, cioè a interrompere un legame con un attore se è possibile raggiungerlo attraverso almeno un altro attore cui si è legati e che è legato ad esso (a differenza della transitività, in questo caso il numero dei legami indiretti è irrilevante, l'importante è la possibilità in sé di raggiungere indirettamente l'altro);
- f. popolarità degli altri: tendenza a creare legami con attori cui tanti membri del network sono legati, cioè con attori che hanno un *indegree* alto;
- g. attività degli altri: inclinazione a creare legami con attori legati a tanti membri del network, cioè con attori che hanno un *outdegree* alto;
- h. chiusura circolare: propensione a creare legami che chiudano cerchi di tre attori e permettano di evitare rapporti gerarchici;
- i. *betweenness*: scelta di creare legami che diano la possibilità di fare da intermediario tra attori che non sono collegati tra loro.

Transitività, equilibrio e attori a distanza due rappresentano tre modi leggermente diversi di misurare la chiusura strutturale, i primi due esprimono direttamente l'idea, il terzo è un indicatore negativo di tale chiusura. I tre effetti hanno una forte covarianza e per non rendere il modello ridondante, si può inserire quello che meglio rappresenta il meccanismo di chiusura nel gruppo osservato, testandoli alternativamente⁵.

⁵ Sulla base della propria esperienza, Snijders afferma che i legami indiretti spesso sembrano essere la misura più efficace per la chiusura strutturale.

Alle forze endogene del network si possono aggiungere effetti – sia semplici sia composti – collegati alle caratteristiche degli attori. I più comuni sono i seguenti:

- l. similarità: inclinazione a creare legami con attori simili a noi, cioè con attori con cui si ha un attributo individuale in comune (ad esempio il titolo di studio), si tratta della tendenza alla selezione omofila;
- m. similarità per reciprocità: scelta di reciprocare legami con attori simili a noi, questo effetto composto unisce una forza collegata alle caratteristiche dell'attore e una forza strutturale, in pratica si misura l'importanza dell'omofilia per la tendenza alla reciprocazione;
- n. caratteristica degli altri: scelta di creare o interrompere legami con i membri del network in base ad un loro attributo individuale, si tratta dell'effetto delle caratteristiche degli attori sulla loro popolarità (*indegree*);
- o. caratteristica di ego: scelta di creare o interrompere legami con i membri del network in base ad un proprio attributo individuale, si tratta dell'effetto delle caratteristiche degli attori sulla loro attività (*outdegree*).

Le caratteristiche degli attori possono essere attributi individuali, come l'età o il titolo di studio, oppure relazioni – sia simmetriche sia asimmetriche – tra le singole coppie ordinate di attori, ad esempio la distanza chilometrica o il rapporto gerarchico all'interno di un'azienda. Per gli attributi relazionali è possibile misurare l'importanza dell'effetto della forza di una relazione per la probabilità di un legame arbitrario e per la reciprocazione.

Se ad un certo punto l'attore i ha la possibilità di cambiare una relazione all'interno del network, valuta la nuova configurazione e la propria posizione al suo interno dopo aver cambiato la relazione con un generico attore j . L'attore i fa questa valutazione per ogni membro del network e sceglie j e la configurazione derivante dal cambiamento della relazione x_{ij} che gli offrono la soddisfazione maggiore. In altre parole, la funzione di valutazione esprime la distribuzione delle preferenze dell'attore i per le diverse configurazioni possibili del network e l'intento dell'attore è quello di massimizzare tale funzione.

1.2. Funzione obiettivo: funzione di dotazione (*endowment function*)

La funzione di dotazione è la seconda componente della funzione obiettivo ed indica la diversa gratificazione provata dall'attore i per la creazione e per l'interruzione del legame x_{ij} . Si denota con l'espressione

$$g_i(\gamma, x, j). \quad (2)$$

Un determinato effetto può avere un'intensità diversa a seconda che si tratti della creazione di un nuovo legame oppure dell'interruzione di un legame esistente: spesso quest'ultima determina la perdita di tutto ciò che un attore ha investito in termini di tempo ed energie in un rapporto (cfr. van De Bunt 1999). La funzione di valutazione non è in grado di dare conto di questa differenza.

Per comprendere la disparità valutativa tra creazione e interruzione di un legame si può pensare alla reciprocità. Rispetto alla creazione di un nuovo legame arbitrario, la reciprocazione di un legame esistente offre un vantaggio perché è meno rischiosa e le incognite sono minori, tuttavia la reciprocità assume un'importanza ancora maggiore per l'interruzione di un legame esistente. È vero, ad esempio, che un'azienda A può preferire comprare un bene da un'azienda B che è già sua cliente e si è rivelata affidabile e professionale piuttosto che da un'azienda C con cui non ha mai avuto rapporti d'affari, però tutto sommato anche un acquisto da C è pensabile e fino a prova contraria non comporta svantaggi. Avendo due fornitori per uno stesso bene e dovendo disdire uno dei due contratti di fornitura, ad A costerebbe però molto meno interrompere il legame con C che quello con B.

γ è il vettore dei parametri che determinano la funzione di dotazione e rappresenta l'entità della differenza tra la creazione e l'interruzione di un legame. Potenzialmente tale differenza si può calcolare per tutti gli effetti inseriti nella funzione di valutazione, in pratica i casi in cui ha senso applicarla sono i seguenti:

- a. reciprocità;
- b. legami indiretti o attori a distanza due;
- c. similarità.

È impossibile rappresentare perfettamente la struttura preferenziale. Inoltre, sull'utilità futura delle diverse configurazioni del network per l'attore i possono influire fattori esterni, come le scelte degli altri attori. Per questo motivo la funzione obiettivo prevede anche una componente casuale ε_i che rappresenta le forze non modellate e si assume indipendente e identicamente distribuita per tutti gli attori, per tutte le configurazioni possibili del network e per tutti i *microsteps* nel tempo. La rappresentazione matematica completa della nuova situazione è data dall'espressione

$$f_i(\beta, x(i \rightsquigarrow j)) + g_i(\gamma, x, j) + \varepsilon_i(t, x, j), \quad (3)$$

dove $i \rightsquigarrow j$ indica il cambiamento dello stato del legame dell'attore i con l'attore j.

1.3. Funzione di distribuzione nel tempo (rate function)

La funzione di distribuzione nel tempo indica la frequenza con cui l'attore i può cambiare qualcosa nei suoi legami in uscita nei periodi tra i singoli momenti di osservazione e si denota con l'espressione

$$\lambda_i(\rho_m, \alpha, x). \quad (4)$$

La funzione dipende dal tasso generico di cambiamento del periodo e dalle caratteristiche individuali e di rete degli attori. Il tasso di cambiamento è dato dalla media dei legami che gli attori cambiano tra due momenti di osservazione successivi. A questo proposito va detto che i doppi cambiamenti non possono essere registrati, cioè che i legami che nel periodo tra i momenti di osservazione t_{m-1} e t_m vengono prima creati e poi interrotti al momento t_m non risultano e i legami prima interrotti e poi ricreati risultano unicamente presenti.

Nel caso più semplice le caratteristiche individuali non sono prese in considerazione, la funzione di distribuzione nel tempo resta costante nell'arco dell'intero periodo tra due momenti di osservazione successivi, è determinata unicamente dal periodo e diventa

$$\lambda_i(\rho_m) = \rho_m. \quad (5)$$

Come l'effetto di densità nella funzione di valutazione, i tassi generici di cambiamento di tutti i periodi tra i singoli momenti di osservazione rappresentano un'intercetta e vanno sempre inseriti nel modello.

Se le caratteristiche degli attori sono rilevanti, il numero di cambiamenti nei legami varia anche da un attore all'altro. α è il vettore dei parametri che rappresentano tali caratteristiche, che si possono riassumere come segue:

- a. effetto degli attributi individuali di un attore sulla possibilità che esso ha di cambiare i propri legami;
- b. importanza della posizione o situazione di un attore all'interno della rete per la frequenza dei cambiamenti nei suoi legami:
 - popolarità (c'è un nesso tra l'*indegree* degli attori e la frequenza dei cambiamenti nei legami?);
 - attività (*outdegree*);
 - reciprocità (gli attori con tanti legami reciproci cambiano i propri legami più o meno velocemente rispetto a coloro che ne hanno pochi?).

In una classe d'asilo, ad esempio, un bambino che possiede molti giocattoli (caratteristica individuale) può cambiare i suoi compagni di gioco molto più spesso di quanto possa fare un bimbo che non ha alcun giocattolo; anche un bambino che in classe è amico di molti altri bimbi (caratteristica di rete) ha più possibilità di cambiare compagni di gioco.

1.4 .Il modello completo

In sintesi, il modello *actor-oriented* per l'interpretazione dell'evoluzione del network è definito da un lato dalle valutazioni degli attori che hanno la possibilità di cambiare un loro legame all'interno del network e dall'altro dalla frequenza di questi cambiamenti. Il primo aspetto si esprime attraverso la funzione obiettivo (composta di funzione di valutazione, funzione di dotazione ed elemento casuale), il secondo attraverso la funzione di distribuzione nel tempo. La tabella 2 riporta le rispettive formule nella colonna del modello completo.

Tab .2 - Modello completo e ridotto per l'evoluzione del network

	Modello completo	Modello ridotto
Funzione obiettivo	$f_i(\beta, x(i \rightarrow j)) + g_i(\gamma, x, j) + \varepsilon_i(t, x, j)$	$f_i(\beta, x(i \rightarrow j))$
Funzione di distribuzione nel tempo	$\lambda_i(\rho_m, \alpha, x)$	ρ_m

I due parametri stocastici riferiti alla struttura preferenziale degli attori esprimono la soddisfazione per lo stato attuale del network (β) e la gratificazione dovuta ai cambiamenti specifici che hanno portato a tale stato (γ), i due parametri relativi alla frequenza con cui gli attori operano i cambiamenti rappresentano il tasso generico di cambiamento (ρ) e le caratteristiche degli attori (α). Tenendo presente che la componente casuale della funzione obiettivo non si modella poiché esprime la parte residua o non spiegata delle preferenze dell'attore e si assume indipendente sia dal momento di osservazione sia dall'attore, nel caso semplificato in cui gli attributi degli attori non siano rilevanti per la frequenza dei cambiamenti nei legami nell'intervallo tra due momenti di osservazione successivi e non ci sia differenza di utilità tra un legame creato e uno interrotto, la specificazione del modello si limita agli effetti relativi alla soddisfazione degli attori per la struttura del network e al tasso generico di cambiamento. In pratica, la funzione obiettivo si riduce alla sola funzione di valutazione e la funzione di distribuzione nel tempo diventa costante nell'intervallo tra due momenti di osservazione successivi; le rispettive formule sono riportate nella colonna per il modello ridotto della tabella 2.

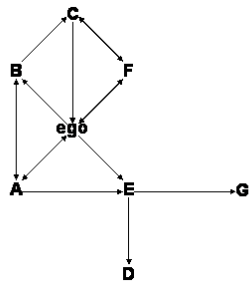
Ad eccezione delle situazioni in cui sproporzioni significative nelle caratteristiche degli attori suggeriscono un'influenza effettiva di tali attributi sulla frequenza dei cambiamenti nei legami è buona norma partire da un modello con la funzione di distribuzione nel tempo costante ed eventualmente complicarlo nel corso dell'analisi. Analogamente, a meno di discrepanze

evidenti tra la valutazione della creazione e dell'interruzione di un legame la funzione di dotazione all'inizio si può considerare nulla, eventualmente in un secondo momento ha senso verificare la differenza di utilità tra legami interrotti e legami creati nel caso di effetti importanti per la specificazione della struttura preferenziale degli attori.

1.5. Esempio esplicativo

Prima di passare all'applicazione pratica del modello, può essere utile ripercorrerne la logica attraverso l'esempio di un piccolo network con otto attori (cfr. Snijders 2005b), riprodotto nella figura 2.

Fig. 2 - Esempio di modello per la dinamica del network



Ego:
 Outdegree = 4
 Legami reciproci = 2
 Legami indiretti = 3

Effetto	Par.	(e.s.)
Tasso cambiam. $t_1 \rightarrow t_2$	3,82	(0,60)
Densità	-1,05	(0,19)
Reciprocità	2,43	(0,40)
Legami indiretti	-0,55	(0,08)

Fonte: adattamento da Snijders (2005b)

Il modello elaborato nell'esempio considera costanti funzione di distribuzione nel tempo e funzione di dotazione e prevede il tasso generico di cambiamento e gli effetti di densità, reciprocità e legami indiretti della struttura preferenziale dell'attore. La relativa funzione di valutazione è

$$f_i(x) = \sum_j (\beta_{dens} x_{ij} + \beta_{rec} x_{ij} x_{ji} - \beta_{dist2} (1 - x_{ij}) \max_h (x_{ih} x_{hj})). \quad (6)$$

β_{dens} , β_{rec} e β_{dist2} sono i parametri dei tre effetti e sono accompagnati dalle rispettive formule:

- x_{ij} esprime l'effetto della densità e vale uno solo se il legame $i \rightarrow j$ esiste, in pratica si calcola l'*outdegree* dell'attore i ;
- $x_{ij} x_{ji}$ esprime l'effetto della reciprocità e vale uno solo se entrambi i legami $i \rightarrow j$ e $j \rightarrow i$ esistono, si contano i legami reciproci dell'attore i ;
- $(1 - x_{ij}) \max_h x_{ih} x_{hj}$ esprime l'effetto delle relazioni indirette, si contano le situazioni in cui $i \rightarrow j = 0$, $i \rightarrow h = 1$ e $h \rightarrow j = 1$, cioè i membri del network da cui l'attore i si trova a distanza due.

Attribuendo ai parametri i valori del modello riportato nella figura 2, la formula per la funzione di valutazione diventa

$$f_i(x) = \sum_j (-1,05x_{ij} + 2,43x_{ij}x_{ji} - 0,55(1-x_{ij})\max_h(x_{ih}x_{hj})). \quad (7)$$

Il segno dei valori dei parametri indica che, oltre ad avere una preferenza per i legami reciproci, gli attori tendono ad avere pochi legami e a non avere legami indiretti.

Il valore della funzione di valutazione è dato dalla somma, ponderata sui parametri, di tutti gli effetti presi in considerazione. Nel caso dell'attore ego, che ha un *outdegree* pari a 4 (A, B, E, F), due legami reciproci (A, F) e tre legami indiretti (C, D, G), si ha

$$f_{ego}(x) = -1,05 \times 4 + 2,43 \times 2 - 0,55 \times 3 = -0,99. \quad (8)$$

La tabella 3 sintetizza numericamente le valutazioni relative ai singoli cambiamenti che l'attore ego può fare e individua nell'interruzione del legame con E il cambiamento più vantaggioso rispetto alla situazione attuale.

Tab. 3 - Vantaggi e svantaggi dei vari cambiamenti nella struttura del network

Cambiamento	Densità	Reciprocità	Legami ind.	Vantaggio
Interruzione con A	3	1	3	-1,38
Interruzione con B	3	2	3	-1,38
Creazione con C	5	3	2	+1,93
Creazione con D	5	2	2	-0,50
Interruzione con E	3	2	1	+2,15
Interruzione con F	3	1	3	-1,38
Creazione con G	5	2	2	-0,50

Fonte: Snijders (2005b)

La nuova funzione di valutazione diventa

$$f_{ego}(x(ego \rightsquigarrow E)) = -1,05 \times 3 + 2,43 \times 2 - 0,55 \times 1 = 1,16. \quad (9)$$

Il vantaggio netto è dato dalla differenza tra nuova e vecchia funzione di valutazione ($1,16 - 0,99 = 2,15$).

In altre parole, riducendo di due unità i legami indiretti (prima erano tre ora è uno solo) e di una unità i legami non reciproci (da quattro diventano tre), ego diminuisce lo svantaggio derivante dalla densità e dagli attori a distanza due e migliora la propria posizione all'interno del network.

Nell'applicazione pratica del modello non è necessario eseguire tutti questi conti e fare i paragoni, calcolo che per tutti gli attori e per ogni ipotetico *microstep* richiederebbe molto tempo. Per individuare le tendenze d'azione degli attori è sufficiente interpretare i parametri degli effetti significativi.

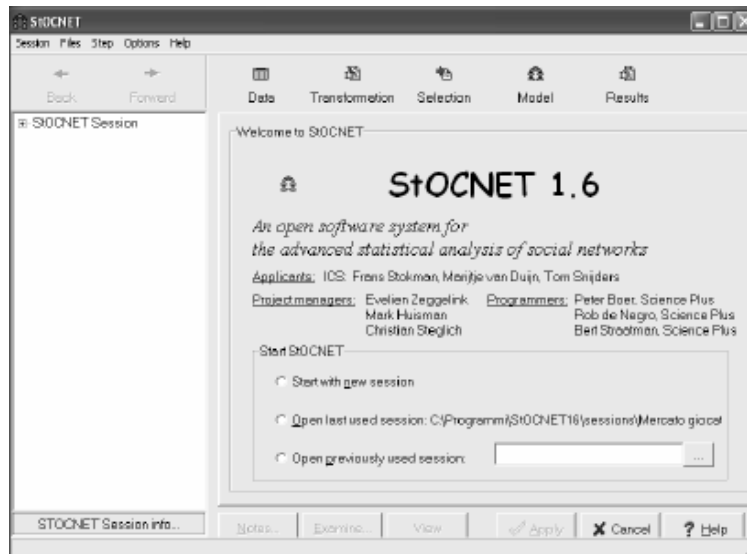
Diversamente da quanto accade per tecniche statistiche come la regressione

lineare o logistica, l'interpretazione dei parametri stimati da SIENA non è immediata, pertanto può risultare difficile, all'inizio, valutare e comparare l'intensità dei diversi effetti inseriti nella funzione obiettivo. Nell'esempio, l'effetto di densità risulta negativo ed effettivamente è abbastanza normale che la creazione di un legame arbitrario abbia più costi che benefici. L'intensità delle altre proprietà strutturali tiene conto della densità: nel caso della reciprocità si registra un valore positivo ($-1,05+2,43=1,38$) e pertanto una preferenza verso la reciprocazione dei legami. All'aumentare della complessità del modello si complica l'intreccio degli effetti e con esso lo schema interpretativo.

2. Applicazione del modello *actor-oriented* in SIENA

Il sistema StOCNET è strutturato in cinque procedure logicamente sequenziali che permettono di inserire i dati, trasformarli, selezionare il subset da analizzare, definire il modello da elaborare e visualizzare i risultati.

Fig. 3 - Maschera iniziale di StOCNET



Fonte: Boer, Huisman, Snijders, Zeggelink (2005)

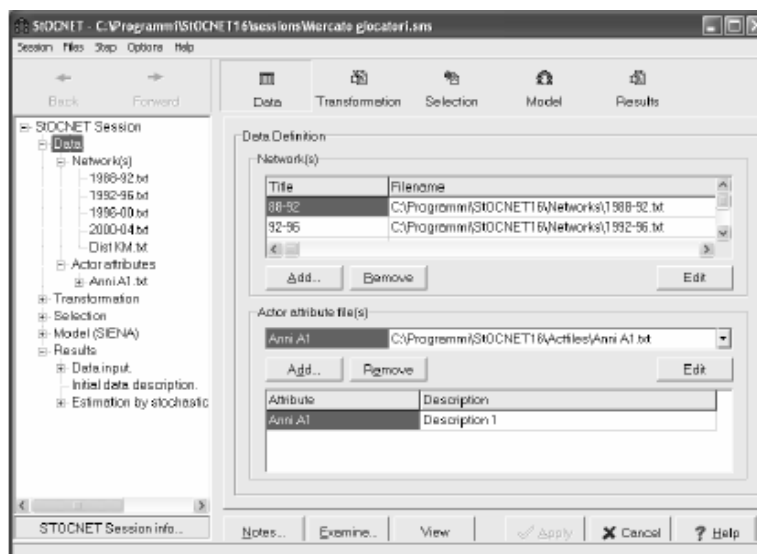
La maschera di apertura di StOCNET consente di scegliere tra l'inserimento di un nuovo set di dati, definito sessione, il recupero dell'ultimo

lavoro fatto e l'apertura di una sessione salvata in precedenza (cfr. fig. 3)⁶.

2.1. Dati, trasformazione e selezione

Indipendentemente dalla voce selezionata, il sistema apre la maschera per la gestione dei dati (cfr. fig. 4).

Fig. 4 - Maschera per la gestione dei dati di StOCNET



Fonte: Boer, Huisman, Snijders, Zeggelink (2005)

Nel caso di un nuovo set di dati, la prima operazione da effettuare è l'inserimento dei file di dati su cui si vogliono eseguire i calcoli. I file devono essere in formato ASCII e si possono importare in StOCNET in formato testo MS-DOS diviso da spazio (non da tabulatore) .txt oppure in formato dati .dat, tenendo presente che i file devono essere privi di riga e colonna di intestazione. È utile rinominare i network e gli attributi inseriti nella sessione di lavoro, poiché StOCNET utilizza nomi come <network1>, <attribute1>, <file1>, che possono creare confusione nella specificazione del modello e rendere meno intuitiva la lettura del file di output.

⁶ Per una trattazione più dettagliata di tutti i procedimenti applicativi si rimanda al manuale di SIENA, scritto da Snijders e tre suoi collaboratori (Snijders, Steglich, Schweinberger, Huisman 2005).

I dati da inserire sono relativi a quattro aspetti del network e dei suoi membri.

- a. Stato del network nei diversi momenti di osservazione: per ogni momento si ha una matrice di adiacenza attore per attore; queste matrici devono essere binarie e le autorelazioni non sono rilevanti, pertanto i valori della diagonale non vengono considerati. Il numero minimo di osservazioni per elaborare un modello evolutivo è due, tuttavia per due sole osservazioni su un network con meno di 30 attori e di 13-15 cambiamenti nei legami non ha comunque senso elaborare un modello con SIENA. Alle volte non si dispone di osservazioni di momenti passati e si costruisce un network retrospettivo, partendo dalle indicazioni sul passato date dai componenti del network attuale (ad esempio, domandando «da quanto tempo sei amico di Tizio?»). Però non è possibile ricostruire la storia degli attori usciti dal network e dei legami interrotti, inoltre questo modo di procedere non è conciliabile con gli assunti della logica dei processi markoviani, poiché viene meno l'influenza del momento t_{m-1} sul momento t_m .
- b. Attributi individuali degli attori, detti covariate individuali: nel caso di un attributo costante nel tempo, come il genere, si ha un file con una riga per ogni attore e una sola colonna valida per tutti i momenti di osservazione; se invece si tratta di attributi soggetti a variazione nel tempo, ad esempio il titolo di studio, si ha una colonna per ogni intervallo tra due momenti successivi di osservazione ($m-1$ colonne per m momenti di osservazione, cfr. fig. 1) e si indica colonna per colonna il valore relativo al momento di osservazione precedente all'intervallo. Per essere modellate correttamente in SIENA, le covariate devono essere ordinali e non semplicemente nominali. Eventuali attributi nominali si possono trasformare dicotomizzandoli oppure creando una variabile *dummy* per ogni categoria. Queste *dummies* assumono valore 1 per gli attori che si trovano nella relativa categoria e 0 per gli attori che si trovano in una qualsiasi delle altre categorie.
- c. Attributi relazionali degli attori, detti covariate diadiche: essendo relazioni, le covariate diadiche si esprimono attraverso una matrice attore per attore, proprio come la relazione di rete oggetto di studio. Gli attributi relazionali possono essere simmetrici, ad esempio la distanza chilometrica, oppure asimmetrici, come il rapporto gerarchico. I valori della matrice devono essere interi e compresi tra 0 e 255. Come per gli attributi individuali, anche per le covariate diadiche è possibile analizzare eventuali variabili nominali grazie all'utilizzo di *dummies* per contraddistinguere coppie di attori che condividono una stessa categoria: basta una *dummy* relazionale se non è rilevante quale categoria condividono due attori, invece bisogna

prevedere una *dummy* relazionale per ogni categoria se si intende specificare di quale categoria condivisa si tratta.

- d. Cambiamenti nella composizione del network: può darsi che alcuni attori siano presenti solo in determinati momenti di osservazione, ad esempio in una classe scolastica nell'arco degli anni escono gli alunni bocciati e magari arrivano nuovi alunni. A seconda della relazione che si studia, può essere più utile considerare gli attori non presenti come righe e colonne di zeri nella matrice di adiacenza (*missing values*⁷) oppure immaginare che si continui a sentire la loro presenza e mantenere righe e colonne dell'ultimo momento di presenza di questi attori (valori 0 e 1 strutturali⁸). In alternativa a questo secondo metodo si possono inserire informazioni sull'eventuale entrata e/o uscita dal network degli attori, contenute in un file che ha una riga per ogni attore e quattro colonne relative rispettivamente (1) all'ultimo momento di assenza, (2) alla frazione di entrata⁹ dell'intervallo successivo a tale momento, (3) all'ultimo momento di presenza, (4) alla frazione di uscita dell'intervallo successivo a tale momento.

Il programma è in grado di gestire un massimo di 500 attori, 99 momenti di osservazione e 10 covariate.

Sui dati inseriti si possono operare alcune trasformazioni: simmetrizzare, ricodificare, dicotomizzare, codificare determinati valori come *missing values*. Si può infine scegliere di lavorare solo su un determinato subset di attori, escludendone alcuni dal modello, sulla base degli attributi, della posizione all'interno della matrice oppure di calcoli specifici.

2.2. Modello

Definiti i file su cui il programma deve eseguire i calcoli e selezionata l'applicazione che si intende usare, cioè SIENA, si specificano i dati e il modello.

Per quanto concerne la specificazione dei dati, SIENA ha bisogno di sapere in che modo combinare e utilizzare i file importati e selezionati: vanno specificate le osservazioni del network (nella giusta sequenza temporale) e le covariate relazionali e individuali (indicando se costanti o variabili) che si

⁷ Se si considerano i legami degli attori non presenti come *missing values* aumenta la frazione di dati mancanti e il network può risultare più sconosciuto che conosciuto.

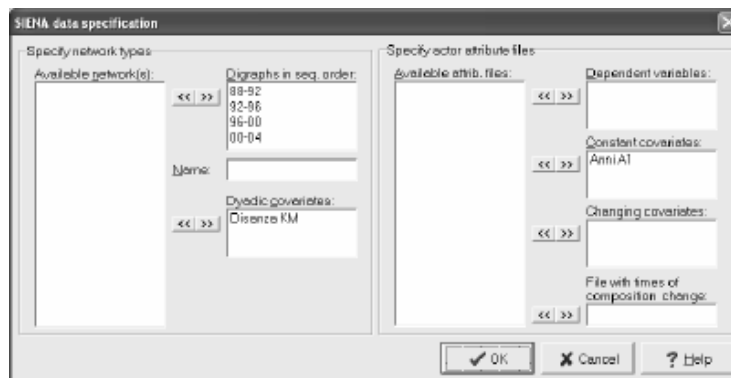
⁸ Un valore è detto strutturale quando non può cambiare. Per 0 e 1 strutturali StOCNET usa rispettivamente i codici 10 e 11.

⁹ Quando non si conosce con esattezza il momento di entrata il valore meno distorto è 0,5. Lo stesso vale anche per la frazione di uscita in colonna (4).

vogliono inserire nel modello.

La figura 5 riproduce la maschera per la specificazione dei dati e si riferisce ad un'analisi dello scambio di giocatori tra le squadre della massima serie della pallamano italiana, osservato negli ultimi quattro quadrienni olimpici, analisi utilizzata a titolo di esempio.

Fig. 5 - Maschera per la specificazione dei dati di StOCNET



Fonte: Boer, Huisman, Snijders, Zeggelink (2005)

Nel caso in cui si studia l'evoluzione del network e si modella il processo di selezione, tutte le caratteristiche degli attori sono variabili indipendenti, mentre la variabile dipendente è data dalla dinamica stessa del network. Se invece si considera anche il processo di influenza e insieme all'evoluzione del network si analizza l'evoluzione del comportamento degli attori, si indica almeno un attributo individuale soggetto a cambiamento nel tempo come variabile dipendente.

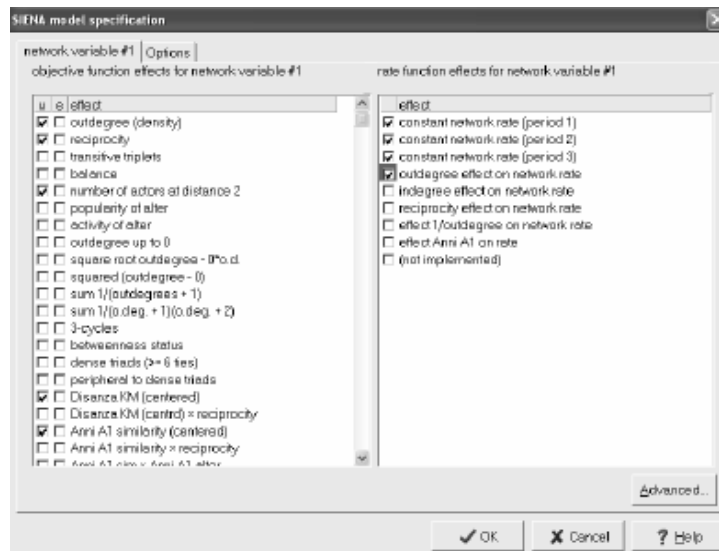
A questo proposito, è importante che prima dell'aggiunta di una variabile dipendente comportamentale si salvi la sessione di lavoro e si proceda ad una nuova sessione, anche semplicemente rinominando quella precedente¹⁰. Le covariate indipendenti non incidono sulla struttura di elaborazione e possono essere sempre inserite nella specificazione iniziale dei dati. Invece quando si hanno variabili comportamentali dipendenti cambia la struttura del modello ed è necessario fare questo doppio passaggio, altrimenti SIENA, sovrascrivendo automaticamente tutti i file che contengono le elaborazioni, perde la memoria dei risultati prodotti.

Definiti i dati, si passa alla specificazione del modello (cfr. fig. 6): bisogna

¹⁰ Le sezioni di lavoro sono salvate come file .sns nella cartella <sessions>.

indicare quali effetti strutturali si vogliono prendere in considerazione per il calcolo. È questa la fase in cui concretamente si decide come definire il modello *actor-oriented* attraverso il quale si vuole interpretare l'evoluzione del network (ed eventualmente l'evoluzione del comportamento all'interno del network). Nella parte sinistra della schermata si modellano i cambiamenti nella struttura relazionale, cioè la funzione obiettivo: la prima colonna è relativa alla funzione di valutazione e la seconda alla funzione di dotazione. La parte destra si riferisce alla frequenza di tali cambiamenti, ossia alla funzione di distribuzione nel tempo.

Fig. 6 - Maschera per la specificazione del modello in SIENA, evoluzione del network



Fonte: Boer, Huisman, Snijders, Zeggelink (2005)

È consigliabile partire con un modello semplice e considerare gli effetti *outdegree (density)* e *reciprocity*, insieme ad uno dei tre effetti di chiusura (*transitive triplets*, *balance* o *number of actors at distance 2*), supponendo nulla la funzione di dotazione. Il modello si può complicare aggiungendo progressivamente effetti di gerarchia (*betweenness* e *3-cycles*) e della posizione degli altri attori (*popularity of alter* e *activity of alter*¹¹). Il programma propone ulteriori effetti, complessi e composti, che però rischiano di appesantire il modello e richiedono un'elaborazione abbastanza lunga. Da

¹¹ Essendo instabili, questi due effetti non vanno inseriti insieme nel modello.

un punto di vista strutturale reciprocità, chiusura e gerarchia solitamente sono sufficienti e permettono di approssimare con precisione apprezzabile le forze endogene che sottendono l'evoluzione del network.

Oltre agli effetti strutturali si possono modellare effetti relativi alle covariate. Per gli attributi individuali solitamente si esplorano *similarity*, *<attribute> alter* e/o *<attribute> ego*, però anche in questo caso sono disponibili effetti composti. Per le covariate diadiche il programma propone un effetto semplice (*centered*¹²) e uno composto (*centered x reciprocity*), nel primo caso si valuta l'importanza della covariata per la probabilità del legame, nel secondo per la reciprocazione dello stesso.

La colonna destra della maschera per la specializzazione del modello è relativa alla funzione di distribuzione nel tempo. Salvo casi in cui si ipotizza a priori una differenza della frequenza dei cambiamenti tra gli attori, si inseriscono solo i tassi generici di cambiamento dei singoli periodi, i cosiddetti *rate parameters*.

Se nella specificazione dei dati è indicata anche una variabile comportamentale dipendente è possibile accedere ad un'apposita schermata per selezionare gli effetti relativi a tale variabile e al processo di selezione. Tra gli effetti della variabile dipendente va sempre preso in considerazione *tendency*, una sorta di intercetta che rappresenta la propensione generale al cambiamento comportamentale. In appendice 2 è riportata una tavola che spiega i principali effetti previsti dal programma per modellare l'evoluzione del comportamento all'interno del network.

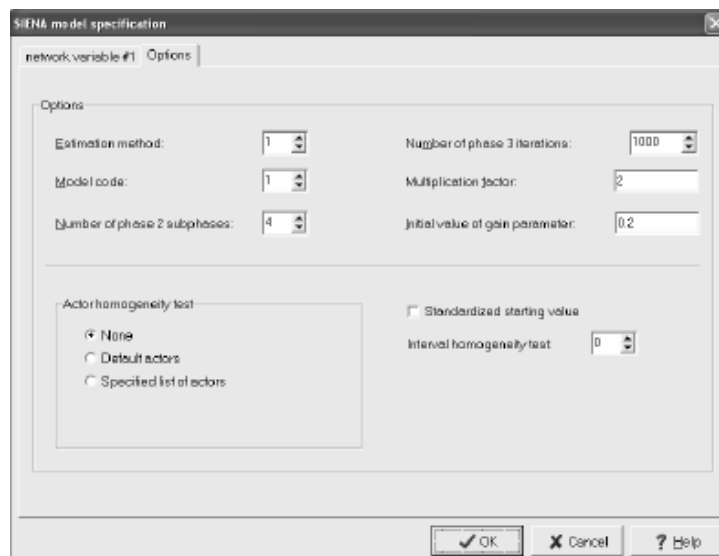
Quando dall'elaborazione di un modello risulta che un parametro è particolarmente alto (anche 100) e gli altri sono molto bassi ed hanno un errore standard piccolo è sensato considerare costante il parametro in questione fissandolo ad un valore alto e modellare gli altri. È quello che succede, ad esempio, in un network in cui i legami non si possono interrompere e c'è un effetto di densità molto forte. I parametri si possono fissare nella sezione *<advanced>* della maschera per la specificazione del modello.

La specificazione del modello prevede un'apposita schermata per selezionare alcune opzioni (cfr. fig. 7) che modificano i criteri che il programma imposta per *default* sulla base dei dati e dell'applicazione selezionata. Le opzioni sono relative alle operazioni di calcolo; le più importanti sono trattate nel dettaglio, dopo una breve descrizione dell'aspetto tecnico-matematico dell'elaborazione del modello.

¹² L'espressione *centered* indica che la variabile è centrata: dal valore effettivo dell'attributo è sottratta la media, in modo da semplificare i calcoli operando con le deviazioni dalla media.

Poiché il modello *actor-oriented* è complesso e lo spazio delle evoluzioni possibili molto ampio, non è realistico effettuare calcoli precisi dei valori attesi e delle probabilità: si applica il metodo dei momenti per individuare i valori attesi che massimizzano la somiglianza con i dati osservati e si approssima la soluzione dell'equazione dei momenti attraverso un algoritmo iterativo progressivo¹³. Tale procedura di simulazione segue tre fasi. In estrema sintesi¹⁴, si parte dai valori osservati dei parametri degli effetti di *default* e si considerano nulli i parametri degli altri effetti selezionati. Su questa base parametrica il programma simula tante evoluzioni casuali e genera così un ampio campione di evoluzioni, effettuando una ricerca casuale di parametri che potrebbero avvicinarsi all'evoluzione osservata. Infine, si valuta la distanza di questo campione dall'evoluzione effettiva (calcolo della convergenza). Sulla base dei risultati ottenuti si può decidere di modificare il modello aggiungendo e/o togliendo alcuni effetti. Il procedimento di stima è lo stesso e i risultati ottenuti nel modello precedente diventano i parametri di partenza.

Fig. 7 - Maschera per le opzioni della specificazione del modello in SIENA



Fonte: Boer, Huisman, Snijders, Zeggelink (2005)

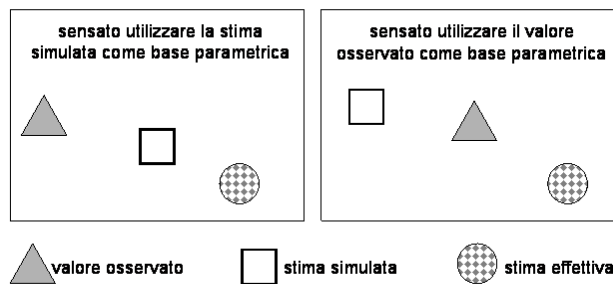
¹³ Il programma segue una procedura di approssimazione stocastica derivata dall'algoritmo di Robbins-Monro (Robbins, Monro 1951).

¹⁴ Per una trattazione dettagliata della procedura si rimanda a Snijders, van Duijn (1997), Snijders (2001) e Snijders (2005a, 2005b).

Per velocizzare la procedura si possono ridurre da 4 a 3 oppure 2 le sottofasi della seconda delle tre fasi dell'elaborazione. La diminuzione delle sottofasi comporta tuttavia una perdita di precisione. Sempre per ridurre i tempi dell'elaborazione e sempre a scapito della precisione, si può diminuire il valore delle iterazioni della terza fase, passando anche da 500 a 100. Il valore iniziale del *gain parameter* esprime la velocità con cui la simulazione si avvicina alla realtà osservata. Abbassandolo si hanno *steps* o salti più piccoli. Mentre ha senso eseguire le operazioni di riduzione delle sottofasi e delle iterazioni a scopo esplorativo, quando cioè si è lontani da un modello interessante, l'abbassamento del valore iniziale del *gain parameter* è utile quando si è vicini ad un risultato interessante, perché riduce le differenze tra una simulazione e l'altra. 0,01 è comunque il valore minimo.

I valori iniziali della prima elaborazione sono i dati osservati, per le elaborazioni successive il programma parte dai valori stimati nell'ultima simulazione, basandosi sul presupposto che la stima simulata sia più vicina alla stima effettiva rispetto ai valori di partenza. Solitamente questo aggiornamento dei valori iniziali è vantaggioso. Quando però le stime simulate non sono ragionevoli, può essere più utile tornare ai valori osservati, selezionando l'opzione di standardizzazione dei valori iniziali. Questa ipotesi è rappresentata graficamente nella figura 8. Se la simulazione successiva elabora stime ragionevoli, è sufficiente deselezionare tale opzione e utilizzare le ultime stime simulate come nuovi valori di partenza.

Fig. 8 - Rapporto tra stima effettiva, stima simulata e valore osservato



Può capitare, ad esempio, che si abbiano quattro parametri significativi ma si decida di eliminare uno degli effetti corrispondenti per esplorare nuove configurazioni del modello, aggiungendo altri effetti. I tre parametri rimasti potrebbero non avere senso senza il parametro eliminato. In questo caso mantenendo i valori ottenuti la stima peggiorerebbe. Ci si accorge di una situazione di questo tipo dall'incapacità del programma di calcolare gli errori

standard o da errori standard altissimi.

Quando i dati relativi all'evoluzione del network sono simmetrici e quindi non diretti, bisogna indicare al programma come trattare le relazioni, scegliendo uno dei cinque modelli proposti. Partendo da una stessa funzione obiettivo, ognuna di queste modalità determina risultati differenti. Le due possibilità più intuitive sono il modello forzato e quello ad iniziativa unilaterale e conferma reciproca. Nel primo caso un attore può decidere unilateralmente di creare o interrompere un legame; in pratica si considerano i dati come se fossero diretti. Nel secondo modello un attore può decidere unilateralmente di interrompere un legame mentre per la creazione di un legame l'altro attore deve essere d'accordo; è un po' quello che succede in tutti i rapporti interpersonali che non sono sanciti da accordi ufficiali o contratti, basti pensare alle relazioni sentimentali. Il secondo modello tiene conto delle funzioni obiettivo di entrambi gli attori coinvolti nella creazione di un nuovo legame, pertanto è più stabile rispetto al primo modello ed è utilizzato come *default* se non ci sono variabili dipendenti aggiunte.

Lanciato il procedimento, il programma inizia l'elaborazione del modello, prendendo in considerazione gli effetti selezionati. Durante il calcolo sono visualizzate le varie fasi ed è possibile una prima verifica euristica (cfr. fig. 9).

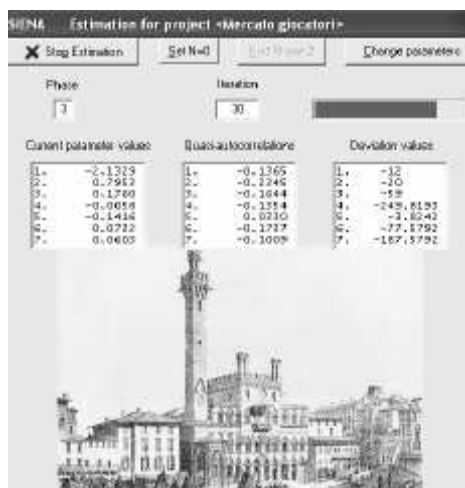
- a. La colonna sinistra riporta le stime dei parametri nei diversi momenti dell'elaborazione; un valore completamente differente dagli altri potrebbe suggerire qualche problema.
- b. La colonna centrale si riferisce alle autocorrelazioni dei singoli parametri; sono auspicabili valori negativi.
- c. I valori della colonna destra indicano deviazioni e si dovrebbero annullare a vicenda, valori solo negativi oppure solo positivi rappresentano un problema.

Procedendo per simulazione, il programma elabora network casuali che rispondano a determinati criteri. Per ogni simulazione si ottengono risultati leggermente differenti, in quanto il caso determina configurazioni sempre diverse. Però a volte può essere utile riprodurre esattamente i risultati di una simulazione. Per farlo bisogna copiare il numero indicato nel file di output come *random state* ed inserirlo nel file *.mo*¹⁵ alla voce *random seed*, impostata per *default* a 0. Inoltre bisogna riportare i parametri ai valori

¹⁵ Questo file è prodotto dal programma durante l'elaborazione, memorizza i dettagli delle operazioni eseguite ed è salvato insieme agli altri file di sistema nella cartella <sessions>. Il file *.out* è l'output presentato dal programma. Il file *.log* contiene informazioni aggiuntive, che possono essere importanti ma appesantirebbero l'output, ad esempio la lista degli effetti possibili e le statistiche che li calcolano.

antecedenti all'elaborazione che si vuole replicare: se è stata fatta una sola simulazione è sufficiente standardizzare i valori iniziali, se le elaborazioni sono più numerose vanno inseriti nel file .mo i valori più recenti dei parametri.

Fig. 9 - Elaborazione del modello actor-oriented in SIENA



Fonte: Boer, Huisman, Snijders, Zeggelink (2005)

Prima di passare ai risultati concreti dell'elaborazione del modello è utile ricordare che SIENA ha un aspetto multilivello per analizzare e confrontare la dinamica di diversi network, ad esempio di diverse classi scolastiche. Si stima separatamente l'evoluzione dei singoli network utilizzando lo stesso modello e poi si combinano i risultati in una metanalisi che permette il confronto. In estrema sintesi, si ha una serie di network cui si applica la stessa definizione del modello, cioè si considerano gli stessi effetti ottenendo parametri ovviamente diversi. Se i network sono pochi, più che l'intensità dell'*effect size* è importante verificare la presenza in sé dell'effetto.

2.3. Risultati

Terminate le procedure di calcolo il programma produce un file di output, le cui informazioni essenziali sono i valori iniziali e finali dei parametri, la convergenza e la matrice di covarianza. Inoltre ci sono alcune statistiche descrittive in merito ai dati inseriti, che possono essere utili per valutare il

network analizzato. In estrema sintesi per essere considerato “buono”, il modello finale deve essere convergente e contenere effetti statisticamente significativi. In appendice 1 è riportato a titolo esemplificativo il file di output relativo all’analisi dello scambio di giocatori nella pallamano italiana, di seguito sono commentati gli aspetti essenziali.

- a. *Degree distributions: out- e/o indegrees* fortemente asimmetrici tra i diversi momenti di osservazione indicano una grande sproporzione nell’attività e/o nella popolarità degli attori. Questa variabilità si potrebbe spiegare introducendo una covariata, ad esempio l’età per un network relativo alla condivisione di anni scolastici. Nell’esempio non si osservano sproporzioni degne di nota.
- b. *Tie changes*: tra un’osservazione e l’altra del network deve esserci continuità, cioè i cambiamenti non devono essere più della metà, in caso contrario bisogna inserire un’osservazione intermedia. I cambiamenti però non devono neppure essere troppo pochi, altrimenti due momenti di osservazione distinti non sono giustificati. Il *turnover* effettivo è dato dalla somma delle colonne $0 \rightarrow 1$ e $1 \rightarrow 0$ della statistica sui cambiamenti, cioè dai legami che si creano e da quelli che si interrompono. Nell’esempio i cambiamenti tra due momenti di osservazione successivi sono relativamente pochi, tuttavia la periodicità dei quadrienni olimpici giustifica un’osservazione ogni quattro anni.
- c. *Initial parameter values*: i valori iniziali dei parametri possono essere causa di eventuali problemi di convergenza alla fine dell’elaborazione. È utile dare un’occhiata a tali valori e fare confronti tra elaborazioni successive.
- d. *Information for convergence diagnosis*: la convergenza tra modello e dati osservati esprime la bontà del modello. La prima colonna della tabella di convergenza riporta la media della distanza dei valori stimati dai dati osservati. L’ultima colonna indica le *t-statistics* per queste deviazioni¹⁶. Trattandosi di convergenze, tutte le *t-statistics* dovrebbero essere inferiori a 0,2 in valore assoluto. Nell’esempio del mercato dei giocatori la convergenza tra dati osservati e simulazione è buona, solo due valori sono di poco superiori a 0,1 in valore assoluto. Alle volte si possono avere *t-statistics* troppo alte nella diagnosi della convergenza anche quando si ha un modello complesso abbastanza vicino ai dati osservati. In tal caso si può diminuire il *gain parameter* oppure rilanciare la simulazione ed attendersi

¹⁶ Su indicazione di Snijders, per poter essere pubblicato, il modello dovrebbe presentare *t-statistics* inferiori a 0,1 in valore assoluto. Orientativamente valori assoluti inferiori a 0,3 sono indicativi di un modello migliorabile, valori superiori suggeriscono l’opportunità di cambiare modello.

t-statistics migliori, partendo dall'idea che l'algoritmo non è perfetto e che si tratta comunque di un processo casuale.

- e. *Estimates and standard errors*: questa sezione dell'output riporta il modello stimato. Ad un livello di significatività di 0,05, l'errore standard deve essere inferiore alla metà del parametro (cioè parametro/errore standard ≥ 2). Le stime per il modello elaborato per lo scambio di giocatori nella pallamano italiana sono riportate nella figura 10. L'idea è che il passaggio di giocatori da una squadra all'altra comporti da un lato uno squilibrio nell'organico e dall'altro un passaggio di informazioni. Entrambe le circostanze possono essere considerate negative per il rendimento agonistico della squadra che cede i propri atleti¹⁷.

Fig. 10 - Stime dei parametri per il modello elaborato per lo scambio di giocatori

0.1	Rate parameter cond. variable period 1	3.1108	(0.4640)
0.2	Rate parameter cond. variable period 2	3.1495	(0.3775)
0.3	Rate parameter cond. variable period 3	3.1203	(0.3559)
Other parameters:			
1. u:	outdegree (density)	-2.1329	(0.0559)
2. u:	reciprocity	0.7953	(0.1765)
3. u:	transitive triplets	0.1760	(0.0513)
4. u:	Distanza KM (centered)	-0.0058	(0.0011)
5. u:	Anni A1 similarity (centered)	-0.1416	(0.1156)
6. u:	Anni A1 ego	0.0732	(0.0092)
7. u:	Anni A1 alter	0.0603	(0.0091)

Fonte: elaborazione statistica con SIENA

Il tasso generico di cambiamento si interpreta intuitivamente: il numero medio di cambiamenti nelle cessioni di giocatori ad altre squadre è di poco superiore a tre nei diversi periodi. Questo denota una certa stabilità nelle squadre ed indica che nell'arco del tempo le strategie di mercato non cambiano molto. La stabilità nella rosa delle squadre si evince anche dal valore negativo dell'effetto di densità: cedere giocatori ad altre squadre può avere costi in termini di coesione e di informazione, quindi di risultato. Cedere atleti a squadre da cui sono stati anche acquisiti giocatori è meno costoso, tuttavia il vantaggio derivante dalla reciprocazione di un legame e delle informazioni non basta a coprire i costi della cessione (la somma tra *density* e *reciprocity* è negativa). L'effetto della transitività è molto debole, in altre parole cedere atleti a compagni che hanno acquisito giocatori da squadre le quali hanno a loro volta acquisito atleti da noi può dare qualche vantaggio in termini di rapporti con le altre squadre, però le cessioni

¹⁷ L'argomento è trattato più dettagliatamente in Savoia (2005).

restano costose (la somma tra *density* e *transitive triplets* è negativa). Dedurre l'assenza di chiusura strutturale dall'effetto debole della transitività sarebbe affrettato, potrebbe essere utile modificare il modello cambiando l'effetto di chiusura.

Per quanto riguarda gli attributi degli attori e gli effetti di selezione ad essi connessi, la distanza chilometrica ha un peso minimo sulle cessioni di giocatori, circostanza plausibile in un contesto sportivo professionistico. Il numero di campionati di massima serie disputati dalle squadre esprime l'istituzionalizzazione e il valore sportivo. Dal modello elaborato risulta che la tradizione agonistica della squadra che cede ed acquisisce un giocatore ha un'importanza minima, anche se una tradizione agonistica consolidata rappresenta un leggero incentivo alla cessione e all'acquisizione. Non è rilevante che le due squadre abbiano una tradizione agonistica simile. Il parametro relativo alla similarità istituzionale delle due squadre non è statisticamente significativo e andrebbe tolto nella successiva elaborazione del modello.

- f. *Covariance matrix*: in fondo al file di output è riportata la matrice (parametro per parametro) di covarianza. I valori sotto la diagonale esprimono la correlazione tra i parametri stimati. Rispetto ad una correlazione classica i valori sono alti: poiché si misura lo stesso network è normale che gli effetti tendano a sovrapporsi. Un valore di 0,94 si può ritenere accettabile, mentre 0,999 è troppo alto e indica che gli effetti si sovrappongono troppo. Le covarianze dell'esempio sono tutte accettabili.

3. Conclusione

Il modello per l'analisi della dinamica del network utilizzato in SIENA permette di interpretare le trasformazioni delle relazioni di rete nel tempo come risultato dell'azione degli attori, i quali scelgono di creare e interrompere legami con l'intento di migliorare la propria posizione all'interno del network. Le valutazioni degli attori sui legami da creare o interrompere dipendono essenzialmente da considerazioni relative alla soddisfazione e alla gratificazione connesse alla struttura relazionale. Si modella l'evoluzione del network come funzione delle scelte relazionali degli attori, determinando la probabilità con cui queste scelte si verificano ed individuando la struttura preferenziale che ne stabilisce il contenuto, cioè specificando quando e cosa cambia all'interno del network. Oltre alle valutazioni preferenziali degli attori il modello considera la frequenza con cui essi operano cambiamenti nella struttura relazionale. Tale frequenza è

espressa dalla funzione di distribuzione nel tempo e dipende dal tasso generico di cambiamento ed eventualmente dalle caratteristiche degli attori e/o dalla loro posizione all'interno del network, mentre il contenuto delle scelte è definito dalla funzione obiettivo, che prende in considerazione la soddisfazione degli attori per lo stato attuale del network (funzione di valutazione) e la gratificazione derivante dalle scelte specifiche di creazione o interruzione di legami che hanno portato a tale stato (funzione di dotazione).

Lo schema interpretativo si basa su alcuni assunti che gli danno un'impronta utilitaristica e limitano l'applicabilità del modello alle situazioni che non entrano in contrasto con queste ipotesi di fondo, esponendolo ad una serie di critiche teoriche (cfr. Steglich, Snijders, Pearson 2004). Si tratta di un tipico modello di utilità *as if* (cfr. Friedman 1953) che, conscio dei propri limiti, offre la possibilità di studiare le complesse dinamiche evolutive delle reti di relazioni sociali partendo dai soggetti che lo compongono. Snijders stesso e i suoi collaboratori sottolineano questo aspetto del loro modello *actor-oriented*. La metodologia su cui SIENA si poggia e il programma stesso sono in continuo e pieno sviluppo e il futuro riserverà sicuramente nuove prospettive che permetteranno di avvicinare sempre più il modello alla concreta evoluzione del network e del comportamento dei suoi membri.

Bibliografia

- Boer P., Huisman M., Snijders T.A.B., Zeggelink E.P.H. (2005), *StOCNET. An Open Software System for the Advanced Statistical Analysis of Social Networks. Version 1.6*, ProGAMMA / ICS, Groningen, <http://stat.gamma.rug.nl/stocnet/> (accesso effettuato il 21.10.2005).
- Friedman M. (1953), *The Methodology of Positive Economics*, in *Essays in Positive Economics*, University of Chicago Press, Chicago, pp. 3-43.
- Holland P., Leinhardt S. (1977a), *A Dynamic Model for Social Networks*, in "Journal of Mathematical Sociology", 5, pp. 5-20.
- Holland P., Leinhardt S. (1977b), *Social Structure as a Network Process*, in "Zeitschrift für Soziologie", 6, pp. 386-402.
- Huisman M., Snijders T.A.B. (2003), *Statistical Analysis of Longitudinal Network Data with Changing Composition*, in "Sociological Methods & Research", 32, pp. 253-287.
- Leenders R.T.A.J. (1995), *Models for Network Dynamics. A Markovian Framework*, in "Journal of Mathematical Sociology", 20, pp. 1-21.
- Leenders R.T.A.J. (1996), *Evolution of Friendship and Best Friendship Choices*, in "Journal of Mathematical Sociology", 21, pp. 133-148.
- McPherson J.M., Smith-Lovin L., Cook J.M. (2001), *Birds of a Feather. Homophily in Social Networks*, in "Annual Review of Sociology", 27, pp. 415-445.
- Norris J.R. (1997), *Markov Chains*, Cambridge University Press, Cambridge.
- Robbins H., Monro S. (1951), *A Stochastic Approximation Method*, in "Annals of Mathematical Statistics", 22, pp. 400-407.

- Savoia L. (2005), *Player Market and Sports Success in the Highest Italian Handball League*, scritto non pubblicato presentato al forum «Social Network Analysis. Advances and Empirical Applications», Oxford.
- Snijders T.A.B. (1996), *Stochastic Actor-oriented Models for Network Change*, in “Journal of Mathematical Sociology”, 21, pp. 149-172.
- Snijders T.A.B. (2001), *The Statistical Evaluation of Social Network Dynamics*, in M.E. Sobel, M.P. Becker (eds.), *Sociological Methodology*, Basil Blackwell, London, pp. 361-395.
- Snijders T.A.B. (2005a), *Models for Longitudinal Network Data*, in P.J. Carrington, J. Scott, S. Wasserman (eds.), *Models and Methods in Social Network Analysis*, Cambridge University Press, New York, pp. 215-247.
- Snijders T.A.B. (2005b), *Simulation-Based Statistical Inference for Evolution of Social Networks*, http://stat.gamma.rug.nl/snijders/siena_s.pdf (accesso effettuato il 21.10.2005).
- Snijders T.A.B., Steglich C., Schweinberger M. (2005), *Modeling the Co-Evolution of Networks and Behavior*, in K. van Montfort, H. Oud, A. Satorra (eds.), *Longitudinal Models in the Behavioral and Related Sciences*, Lawrence Erlbaum, 2006 in stampa.
- Snijders T.A.B., Steglich C., Schweinberger M., Huisman M. (2005), *Manual for SIENA 2.1*, http://stat.gamma.rug.nl/stocnet/downloads/sie_man21b.pdf (accesso effettuato il 21.10.2005).
- Snijders T.A.B., van Duijn M.A.J. (1997), *Simulation for Statistical Inference in Dynamic Network Models*, in R. Conte, R. Hegelsmann, P. Terna (eds.), *Simulating Social Phenomena*, Springer, Berlin, pp. 493-512.
- Steglich C.E.G., Snijders T.A.B., Pearson M. (2004), *Dynamic Networks and Behavior. Separating Selection from Influence*, sottoposto alla pubblicazione, http://stat.gamma.rug.nl/snijders/SSP_total.pdf (accesso effettuato il 29.10.2005).
- Taylor H.M., Karlin S. (1998³), *An Introduction to Stochastic Modeling*, Academic Press, New York.
- Thaler R. (1980), *Toward a Positive Theory of Consumer Choice*, in “Journal of Economic Behavior and Organization”, 1, pp. 39-60.
- van de Bunt G.G. (1999), *Friends by Choice. An Actor-Oriented Statistical Network Model for Friendship Networks through Time*, Thesis Publishers, Amsterdam.
- Wasserman S. (1977) *Stochastic Models for Directed Graphs*, Tesi di dottorato, University of Harvard, Harvard.
- Wasserman S. (1979), *A Stochastic Model for Directed Graphs with Transition Rates Determined by Reciprocity*, in K.F. Schuessler (ed.), *Sociological Methodology 1980*, Jossey-Bass, San Francisco, pp. 392-412.
- Wasserman S. (1980), *Analyzing Social Networks as Stochastic Processes*, in “Journal of the American Statistical Association”, 75, pp. 280-294.

Appendice 1: Esempio di file di output di SIENA

```
*****
          Mercato giocatori.out
*****
Filename is Mercato giocatori.out.

This file contains primary output for SIENA project <<Mercato
giocatori>>.
File Mercato giocatori.log contains secondary output.

Date and time: 29/10/2005 11.49.21

SIENA version 2.2b

@1
Data input.
=====

Read basic information file Mercato giocatori.IN.
4 observations,
38 actors,
1 dependent network variable,
0 dependent actor variables,
1 file with constant actor covariates,
0 exogenous changing actor covariates,
1 constant dyadic covariate,
0 exogenous changing dyadic covariates,
no file with times of composition change.

@2
Reading network variables.
-----

Reading digraph files for the 1st network variable:

File C:\~1988-92.txt contains observation moment 1.
nonzero code is 1; missing codes are ... (none).
File C:\~1992-96.txt contains observation moment 2.
nonzero code is 1; missing codes are ... (none).
File C:\~1996-00.txt contains observation moment 3.
nonzero code is 1; missing codes are ... (none).

File C:\~2000-04.txt contains observation moment 4.
nonzero code is 1; missing codes are ... (none).

For file C:\~1988-92.txt, degree distributions are as follows:
Nodes
  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20
 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38
in-degrees
  0  0  0  0  3  1  1  1  0  1  0  0  0  0  0  0  1  0  0  0
  0  0  3  1  0  5  0  2  0  1  0  0  3  0  2  0  0  3
out-degrees
  0  0  0  0  1  0  0  0  0  1  0  0  1  0  0  1  2  4  0  0
  0  0  3  0  0  2  3  0  0  1  0  0  1  0  3  2  0  3
```

No missing data for observation 1 (file C:\~1988-92.txt).

For file C:\~1992-96.txt, degree distributions are as follows:

```
Nodes
  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20
 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38
in-degrees
  0  1  2  0  5  2  1  3  0  0  0  6  0  1  1  0  3  1  0  4
  2  0  6  1  0  7  0  1  1  0  2  0  3  0  2  1  7  5
out-degrees
  0  0  0  0  5  4  1  5  0  1  0  4  1  2  1  1  4  2  0  0
  3  0  3  2  0  6  0  4  1  0  2  0  2  0  6  1  4  3
```

No missing data for observation 2 (file C:\~1992-96.txt).

For file C:\~1996-00.txt, degree distributions are as follows:

```
Nodes
  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20
 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38
in-degrees
  0  1  1  4  1  5  3  8  0  3  0  1  1  4  0  0  0  1  0  5
  2  3  7  3  1  9  0  0  1  1  2  0  8  0  3  0  3  6
out-degrees
  1  2  3  0  7  5  2  5  1  0  1  4  3  0  1  0  3  1  0  4
  4  1  4  1  0  5  0  0  2  3  1  0  7  0  6  0  5  5
```

No missing data for observation 3 (file C:\~1996-00.txt).

For file C:\~2000-04.txt, degree distributions are as follows:

```
Nodes
  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20
 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38
in-degrees
  4  8  0  7  0  0  1  3  3  0  0  7  1  3  2  0  1  0  3  2
  3  0  0  2  1 10  0  0  1  0  4  4  4  1  1  0  0  5
out-degrees
  0  3  0  9  0  0  3  5  1  1  0  5  2  4  0  0  1  1  2  4
  1  4  6  1  2  6  0  0  1  0  4  0  5  1  4  0  1  4
```

No missing data for observation 4 (file C:\~2000-04.txt).

Name of this network variable:(none given).

```
@2
Reading constant actor covariates.
-----
```

StOCNET error bypassed when reading covariate data file named "C:\~Anni A1.txt"

Covariate data file C:\~Anni A1.txt with 1 variables, named:
ANNI A1 (code for missing data is not given, SIENA will treat scores of -9 as missing data).

A total of 1 non-changing individual covariate.

Number of missing cases:
Anni A1 0 (0.0 %)

Means of covariates:
Anni A1 5.368

The mean value is subtracted from the covariate.

@2

Reading constant dyadic covariates.

Dyadic covariate named DISANZA KM read from file C:\~Dist KM.txt.
No code for missing data is specified; SIENA will treat scores of -9 as
missing data.

A total of 1 dyadic covariate.

Number of tie variables with missing data:

Disanza KM 0 (0.0 %)

Means of covariates:

Disanza KM 56.269

This mean value is subtracted from the dyadic covariate.

For the following statistics, missing values (if any) are not counted.

For the similarity variable calculated from each actor covariate,
the mean is subtracted.

These means are :

Similarity Anni A1 : 0.6480

The mean structural dissimilarity value subtracted in the
balance calculations is 0.078447 .

@1

Initial data description.
=====

For the following statistics, missing values (if any) are not counted.

Network density indicators:

observation time	1	2	3	4
density	0.020	0.048	0.062	0.058
average degree	0.757	1.838	2.351	2.189
number of ties	28	68	87	81
missing fraction	0.000	0.000	0.000	0.000

Tie changes between subsequent observations:

periods	0=>0	0=>1	1=>0	1=>1	Distance	
Missing						
1 ==> 2	1313	65	25	3	90	0 (0%)
2 ==> 3	1268	70	51	17	121	0 (0%)
3 ==> 4	1255	64	70	17	134	0 (0%)

Dyad counts:

observation	total	mutual	asymm.	null
1.	703	1	26	676
2.	703	13	42	648
3.	703	14	59	630

4. 703 9 63 631

Standard values for initial parameter values

```
-----  
constant network rate (period 1)      1.5824  
constant network rate (period 2)      1.8781  
constant network rate (period 3)      1.9801
```

outdegree (density) -1.4014
(weighted average of estimates under the (trivial) independent ties
model;

estimates and weights are as follows:

```
period      1      2      3  
density par.est. -1.470 -1.331 -1.404  
weight      0.323  0.331  0.346 )
```

Model descriptions are written to file Mercato giocatori.log.

Initialisation of project <<Mercato giocatori>> executed succesfully.

```
-----  
New analysis started.  
Date and time: 29/10/2005 11.50.30  
New results follow.  
-----
```

@1

Estimation by stochastic approximation algorithm.

=====
Random state is initialized at 417124650.

Model Type 1: Standard actor-oriented model.
Estimation method: conditional moment estimation.
Conditioning variable is the total number of observed changes
("distance")
in the network variable.
Distances for simulations are
period : 1 2 3
distance : 90 121 134 .
Initial value of gain parameter is 0.2000000 .

Initial parameter values are

```
0.1 Rate parameter period 1      1.5824  
0.2 Rate parameter period 2      1.8781  
0.3 Rate parameter period 3      1.9801  
1. u: outdegree (density)      -1.4014  corresp. to statistic 4  
2. u: reciprocity              0.0000  corresp. to statistic 5  
3. u: transitive triplets      0.0000  corresp. to statistic 6  
4. u: Distanza KM (centered)   0.0000  corresp. to statistic 20  
5. u: Anni A1 similarity (centered) 0.0000  corresp. to statistic 22  
6. u: Anni A1 ego              0.0000  corresp. to statistic 26  
7. u: Anni A1 alter            0.0000  corresp. to statistic 30
```


Observed values of target statistics are

1. Number of ties	236.0000
2. Number of reciprocated ties	72.0000
3. Number of transitive triplets	163.0000
4. Sum of ties × Dianza KM	-3259.8862
5. Amount of similarity on Anni A1	-21.7911
6. Sum of outdegrees × Anni A1	801.0576
7. Sum of indegrees × Anni A1	800.0576

7 parameters, 7 statistics

@2
End of stochastic approximation algorithm, phase 3.

Total of 1859 iterations.
Parameter estimates based on 1359 iterations,
basic rate parameters as well as
convergence diagnostics, covariance and derivative matrices based on 500
iterations.

Information for convergence diagnosis.
Averages, standard deviations, and t statistics for deviations from
targets:

1.	-1.280	14.071	-0.091
2.	-1.036	13.791	-0.075
3.	-5.570	61.652	-0.090
4.	17.341	649.547	0.027
5.	-0.180	5.250	-0.034
6.	-12.100	110.786	-0.109
7.	-13.076	113.890	-0.115

Good convergence is indicated by the t-statistics being close to zero.

@2
Estimation results.

Regular end of estimation algorithm.
Total of 1859 iteration steps.

@3
Estimates and standard errors

0.1	Rate parameter cond. variable period 1	3.1108	(0.4640)
0.2	Rate parameter cond. variable period 2	3.1495	(0.3775)
0.3	Rate parameter cond. variable period 3	3.1203	(0.3559)

Other parameters:

1. u:	outdegree (density)	-2.1329	(0.0559)
2. u:	reciprocity	0.7953	(0.1765)
3. u:	transitive triplets	0.1760	(0.0513)
4. u:	Dianza KM (centered)	-0.0058	(0.0011)
5. u:	Anni A1 similarity (centered)	-0.1416	(0.1156)
6. u:	Anni A1 ego	0.0732	(0.0092)
7. u:	Anni A1 alter	0.0603	(0.0091)

@3

Covariance matrices

Covariance matrix of estimates (correlations below diagonal):



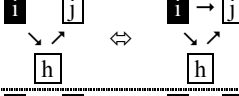
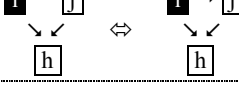
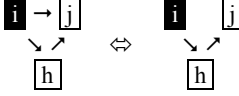
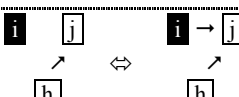
0.003	-0.001	-0.001	-0.000	0.002	-0.000	-0.000
-0.092	0.031	-0.002	0.000	-0.002	-0.001	-0.000
-0.214	-0.258	0.003	0.000	-0.001	-0.000	-0.000
-0.057	0.299	0.223	0.000	-0.000	-0.000	-0.000
0.242	-0.098	-0.220	-0.283	0.013	0.000	0.000
-0.179	-0.317	-0.336	-0.207	0.036	0.000	0.000
-0.203	-0.307	-0.416	-0.248	0.213	0.454	0.000

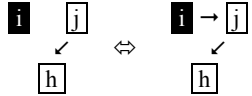
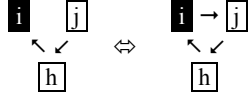
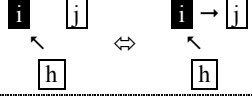




Derivative matrix of expected statistics X by parameters and covariance/correlation matrix of X are written to file Mercato giocatori.log.

Fonte: elaborazione statistica con SIENA adattata



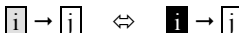


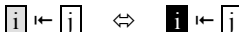
Appendice 2: Effetti per modellare l'evoluzione del network e del comportamento

Evoluzione del network

Effetto	Statistica	Descrizione verbale	Spiegazione grafica
<i>Density</i> (outdegree)	x_{ij}	Scelta di creare legami arbitrari con un qualsiasi altro membro del network	
<i>Reciprocity</i>	$x_{ij}x_{ji}$	Scelta di reciprocare legami	
<i>Transitive triplets</i> (effetto di chiusura strutturale)	$x_{ij} \sum_h x_{ih}x_{hj}$	Scelta di "fare amicizia con gli amici degli amici". Tendenza di i a creare un legame con j se i ha un legame con h che a sua volta è legato a j	
<i>Balance</i> (effetto di chiusura strutturale)	$x_{ij}structsim_{ij}$	Scelta di creare legami con attori strutturalmente simili (equivalenza strutturale), cioè con attori con cui si condividono altri legami.	
<i>Number of actors at distance 2</i> (effetto di chiusura strutturale)	$\begin{cases} 1 \text{ if } between(h,i;j) = 1 \text{ for some } h \\ 0 \text{ else} \end{cases}$	Scelta di avere uno o più intermediari nel rapporto con un altro attore (in <i>transitive triplets</i> l'intermediario è uno). Tendenza di i a interrompere il legame con j se i ha un legame con h che è a sua volta legato a j. Indicatore negativo della chiusura	
<i>Popularity of alter</i>	$x_{ij} \sum_h x_{hj}$	Scelta di creare legami con attori che hanno un alto <i>indegree</i>	

<i>Activity of alter</i>	$x_{ij} \sum_h x_{jh}$	Scelta di creare legami con attori che hanno un alto <i>outdegree</i>	
<i>3-cycles</i>	$x_{ij} \sum_h x_{jh} x_{hi}$	Scelta di creare legami che chiudano cerchi di tre elementi. Indicatore negativo della struttura gerarchica	
<i>Betweenness</i>	$\sum_h \text{between}(i; hj)$	Scelta di creare legami che diano la possibilità di fare da intermediario tra persone che non sono in contatto diretto	
<i>Similarity</i> (effetto relativo a una caratteristica degli attori)	$x_{ij} sim_{ij}$	Scelta di creare legami con attori simili. Tendenza di i a scegliere attori j con cui ha una caratteristica individuale in comune (omofilia, selezione)	
<i>Similarity x reciprocity</i> (effetto relativo a car. attori)	$x_{ij} x_{ji} sim_{ij}$	Scelta di reciprocare legami con attori simili. Importanza della tendenza all'omofilia per la tendenza alla reciprocazione: interazione di due effetti	
<i>Behavior of alter</i> (effetto relativo a car. attori)	$x_{ij} z_j$	Scelta di creare o interrompere legami in base alle caratteristiche degli altri attori. Effetto delle caratteristiche dell'altro sulle scelte di i	
<i>Behavior of ego</i> (effetto relativo a car. attori.)	$x_{ij} z_i$	Scelta di creare o interrompere legami in base alle proprie caratteristiche. Effetto delle caratteristiche di i sulle sue scelte	

Evoluzione del comportamento

Effetto	Statistica	Descrizione verbale	Spiegazione grafica
<i>Tendency</i>	z_i	Tendenza comportamentale principale	
<i>Indegree</i>	$z_i \sum_j x_{ji}$	Effetto della propria popolarità sul comportamento	
<i>Outdegree</i>	$z_i \sum_j x_{ij}$	Effetto della propria attività sul comportamento	
<i>Similarity</i>	$\sum_j x_{ij} sim_{ij}$	Effetto dei legami sul comportamento. Si tratta della tendenza ad assimilare il proprio comportamento a quello degli attori cui si è legati (contagio, influenza)	
<i>Similarity x reciprocity</i>	$\sum_j x_{ij} x_{ji} sim_{ij}$	Effetto dei legami reciproci sul comportamento. Si tratta della tendenza ad assimilare il proprio comportamento a quello degli attori con cui si ha un legame reciproco	
<i>Isolation</i>	$z_i \cdot isolate(i)$	Effetto dell'isolamento all'interno del network sul comportamento	

Fonte: adattamento da Steglich, Snijders, Pearson (2004)